DOCUMENT RESUME

ED 197 231 CG 014 878

AUTHOR Dush, David M.; Kaoukis, George
TITLE Measuring Affect from Speech: The Mahl Speech Disturbance Indices.
PUB DATE Sep 80
NOTE 19p.; Paper presented at the Annual Convention of the American Psychological Association (88th, Montreal, Quebec, Canada, September 1-5, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Affective Measures; *Anxiety; Correlation; Individual Characteristics; Personality Measures; *Personality Traits; *Predictive Validity; Predictor Variables; *Speech Habits; *Test Reliability; Verbal Communication
IDENTIFIERS *Mahl Speech Disturbance Indices

ABSTRACT

The measurement of personality characteristics via the analysis of verbal communication is of interest due to its unobtrusive, accessible, and objective nature. The Mahl Speech Disturbance Indices, a procedure for assessing anxiety on the basis of speech disturbance patterns, was examined to assess the nature and reliability of the indices and their validity as measures of anxiety. A battery of trait anxiety, state affect, and social desirability measures was administered to 96 college students in the context of an intelligence testing experiment. The Mahl indices were derived from audio recordings of speech from 78 subjects. The traditional "Non-Ah" anxiety index failed to correlate significantly with any of the criteria. A two-factor representation of the seven "Non-Ah" ratios yielded a significant canonical correlation with affect measures. Prediction was improved when quadratic and cubic "Ah" components were incorporated. Ultimately, canonical functions emerged which differentially predicted the anxiety, hostility, and depression measures via three linear combinations of the disturbance ratios. (Author/NRB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED197231

Measuring Affect from Speech: The Mahl Speech Disturbance Indices

David M. Dush
Kent State University

George Kaoukis
University of Manitoba

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
David M. Dush

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CG 014878

Presented at the 88th annual convention of the American Psychological Association, Montreal, Quebec, 1980. Direct correspondence to David M. Dush, Department of Psychology, Kent State University, Kent, Ohio, 44242.

Measuring Affect from Speech: The Mahl Speech Disturbance Indices

The measurement of personality characteristics via the analysis of verbal communication is clearly of wide ranging interest for its unobtrusive, accessible, and objective nature. A specific, applied procedure was developed by Mahl (1956) for assessing anxiety on the basis of patterns of speech disturbance. The technique consists of counting the frequency of occurrence of eight categories of speech disturbance, such as stuttering, omitting appropriate words, uttering incoherences, or leaving sentences incomplete. Disturbance ratios were formed by compensating for the total number of words spoken, and these were found to correlate moderately with clinical judgments of anxiousness.

The "Non-Ah" ratio, based on seven of the eight categories, was found to be the best predictor, and subsequently continued to be referenced and employed in clinical research, especially as one of a battery of tests in multi-method measurement (e.g., Farina and Holzberg, 1970; Smith, 1976). However, to date little further attention has been concentrated on the specific psychometric characteristics of the Mahl speech disturbance ratios. Given that such a metric is apparently still in demand, the present study was undertaken to further assess the nature and reliability of the indices, and their validity as measures of anxiety.

A convergent validation approach (Campbell and Fiske, 1959) was presently adopted. The intercorrelations with other measures

of affect are of particular interest given the multivariate context of the Non-Ah's typical recent applications. Also, the unidimensionality implicit in the Non-Ah is of questionable tenability and thus the state-trait and other affective dimensions were held to be necessary considerations.

Method

Subjects were 96 male and female students in introductory psychology courses who earned extra credit for participating. Subjects were run in five randomly assigned groups in a language instruction facility equipped with individual audio recording and monitoring consoles. The experiment, labeled "Ability," was presented to subjects as a study of an intelligence test and the factors affecting performance. The testing context was presented so as to maximize its anxiety arousing stimulus value. The critical points in both written and verbal instructions were that (1) speed and accuracy are essential, (2) the test may be found strenous and difficult, (3) it's one of the most accurate measures of ability available, (4) guessing will lower the score, and (5) the time limit will not allow room for hesitation. Three sample items were included on the face sheet instructions which were intended to be perceived as very difficult by all subjects.

Three audio recordings of speech segments were collected for purposes of deriving the Mahl indices: One at the very beginning of the experiment, one after "anxiety arousing" instructions and procedures in preparation for the intelligence test administration, and one immediately following actual administration of the I.Q.

test. In interest of measuring relatively stable components of anxiety, the Debilitating Anxiety Scale (DAS) (Alpert and Haber, 1960), Test Anxiety Scale (TAS), and General Anxiety Scale (GAS) (Sarason, 1972) were administered following the first recording. A ten item, true-false social desirability scale, and a questionnaire of demographic data were also completed at this time. The more labile affective components were measured, coincident with the first and second speech recordings, by the Anxiety, Hostility, and Depression scales of the Multiple Affect Adjective Checklist (MAACL) (Zuckerman and Lubin, 1965). The IQ test administered, was identified as the "Stanton Intelligence Battery" in a bogus cover sheet, but was actually a modified form of the Wonderlic Personnel Test, Form V (Wonderlic, 1959).

The audio recordings were made from individuals' spontaneous four minute discussions of designated topics. The influence of topical content for the second speech (neutral-topic ("food service at the university") versus anxiety-associated topic ("how you feel about being evaluated")) was evaluated as a two-level factorial component in the design. Group 1, 2, and 5 formed the neutral-topic treatment condition. Procedures were equivalent for all groups except for the topic of the second recording. Most instructions were communicated to subjects from a control room, via headphones at each console. The spontaneous speech and other measures were explained as factors associated with intellectual performance. Subsequent remarks from subjects indicated that the deception had been effective. All subjects were debriefed regarding misleading

aspects of manipulation following the experiment.

Results

All Mahl (audio) data for 18 subjects was unscorable due to equipment malfunction or interfering audio characteristics, such as inaudibly low voice projection or extreme accents. The scored recordings for 35 female and 43 male subjects were associated with inter-rater reliability ranging from .56 ("Omissions") to .98 ("Ahs"), with a median value of .835, based on 10% of the recordings.

The Non-Ah ratio was calculated by dividing cumulative frequency of disturbances (excluding "Ahs") by the total number of words in that segment of speech. For the sum of Non-Ah disturbances, inter-rater reliability was .96, based on the same 10% sample. As indicated in Table 1, the three labile affect measures intercorrelated highly, as did trait measures, but only MAACL Anxiety correlated significantly with trait anxiety measures. However, the traditional Non-Ah ratio failed to correlate significantly with any of the six criterion measures. Curvilinear relationships were not apparent in inspection of scatterplots.

Additional analyses were pursued with individual disturbance ratios, similarly calculated separately for each of the eight disturbance categories. The Non-Ah individual ratios were submitted to a principal components analysis and orthorgonal rotation, producing two factors accounting for 51% of the variability (see Table 2). The two subsequent factor composite scores were correlated simultaneously with the six criterion measures, in a canonical

correlation analysis, and found to produce a canonical correlation ($\underline{R}c$) of .424 ($\underline{p} < .05$). The trait measures contributed very little to this function. A second canonical analysis with the factor composites as predictors and the three state affect criterion measures produced a first canonical function which attained significance ($\underline{R}c = .415$, $\underline{p} < .05$), with virtually no loss in predictive power. Loadings were heaviest and exclusively positive on factor I in the predictor set and MAACL Anxiety in the criterion set. This appeared to confirm measurement of a common trait between the two sets of variables, presumably a state or situational dimension of anxiety.

Insert Tables 1 & 2 about here

Exploration of additional variability accounted for by the Ah ratio in a series of multiple correlations with the Non-Ah ratios suggested a complex interrelationship. Initially, simply including the Ah ratio improved prediction: The first canonical correlation of the Mahl factors and Ahs with the state affect measures rose to .493 ($\underline{p} < .05$). Components of $(Ah)^2$ and $(Ah)^3$ also added substantially to prediction.

In interest of applied practicality and statistical power, it was shown that the influence of the Ah components as predictors could be equitably expressed as a "correction" to the total word value used in calculating the individual Non-Ah ratios, preserving the pattern of loadings as well as the strength of relationship.

Proceeding from exploratory canonical correlation analyses and trials of various combinations of the linear, quadratic, and cubic Ah corrections, the correction of adding $((Ah)^2 - (Ah)^3)$ to the word total was found to produce dramatically superior prediction. The linear Ah component no longer added meaningfully to the above corrected Non-Ah ratios. As apparent in Figures 1 and 2, this correction acts to rapidly reduce the comparative word total as the number of Ahs becomes large, effecting an increase in the magnitude of the disturbance ratios in geometric proportion to the number of Ahs.

Insert Figures 1 & 2 about here

In this final corrected form, the first canonical correlation with MAACL measures rose to .631 ($p < .01$) for the 7 corrected disturbance ratios for the first recording segment. For the second recording segment, the second canonical function also attained significance ($Rc_1 = .610$, $p < .001$; $Rc_2 = .463$, $p = .014$)

Insert Table 3 about here

In further analysis, intended as only suggestive, it was determined that the Ah contribution was maximized when the correction factor was multiplied by 1.2. At this point the correlation for the first recording rose to .710 ($p < .001$), and for the second

recording all 3 canonical functions now attained significance ($\underline{R}c_1$ = .717, $p < .001$; $\underline{R}c_2$ = .443, p = .009; $\underline{R}c_3$ = .391, p = .041). Such fine discriminations of complex curvilinearity place excessive demand on present sample characteristics, however (see Table 3).

In all analysis with corrected ratios the first canonical function loaded definitively on MAACL Anxiety among the criteria, the second canonical function loaded definitively on Hostility, and the third function on Depression. Loadings in the predictor set typically indicated substantial contribution of each of the 7 Non-Ah disturbance ratios, with Stutter and Sentence Incompletion contributing least (see Table 3).

Other variables investigated, such as sex, speech topic, and I.Q. test performance, failed to relate significantly to the Non-Ah ratios. It is paricularly noteworthy that the social desirability index also failed to correlate significantly with the Mahl indices. Yet social desirability did in fact correlate significantly with the DAS and MAACL Anxiety, Hostility, and Depression, which may attest to one of the advantages of the unobtrusiveness of the Mahl measurement.

Discussion

Four major thrusts of the implications of present results can be identified. First, the Non-Ah ratio as originally developed and subsequently used was not found to be predictive of any other affect measures presently employed. Second, multivariate analysis clearly demonstrated the utility of considering the Mahl ratios as multidimensional, and as predictive of multiple dimensions of

labile affect. Third, contrary to traditional use, the Ah Category was found to contribute substantially to prediction. Finally, the optimal predictive function incorporating all 8 disturbance categories was obtained with the inclusion of quadratic and cubic components of the Ah ratio. Albeit additing complexity, the inclusion of curvilinear components, and the full consideration of multivariate relationships, accounted for an improvement in predictability by a factor of about 5:1 over the Non-Ah ratio in its traditional form. At best, in the strongest set of canonical correlations presently found, predictability attained roughly 85% of the total variability.

The findings call into question continued use of the Non-Ah index in an uncorrected and univariate form. Even without correction, factor analysis suggested multidimensionality of the interrelationship of the various disturbance ratios. Moreover, a multivariate consideration of the curvilinear contribution of Ahs not only added greatly to predictive power but may have substantive bearing on interpretation of the relationship of Ahs and the remaining disturbance ratios to affect. Findings suggest an intricate relationship between the prevalence of Ahs -- a relatively common utterance -- and the remaining disturbance categories. In contrast to original findings that Ahs were largely inconsequential (Mahl,1956), the present data suggests that Ahs effect the most salient influence on predictability of all single disturbances. This is manifested as a sharp increase in the value of corrected Mahl disturbance ratios as Ahs increase in relative frequency,

according to a geometric progression.

The intrinsic nature of the relationship of Ahs to affect is highly complex, and not independent of the occurrence of other types of speech disturbances. In one regard the interpretive meaning of the Ah is secondary to its utility in improving prediction: On the other hand, further research to explore the Ah's inter-relationship with affect and speech disturbance is clearly desirable psychometrically, and may be critical to advancing understanding of the interface of speech and affect.

With such complex nonlinearity under consideration, the isolation of the optimal correction factor is difficult and demanding on the present sample. Yet the five-fold increase in predictability lends itself to credibility of the appearance of curvilinearity of the order observed. Further research is indicated, including a cross-validation and replication study to converge on optimal correction and variable loadings for producing composite indices for applied use. While the specific parameters may well vary somewhat in replications, the magnitude of effects presently found argues for thorough and consistent consideration of curvilinearity and multi-dimensionality in further related research.

References

Alpert, R., & Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Farina, A., & Holzberg, J. D. Anxiety level of schizophrenic and control patients and their parents. Journal of Abnormal Psychology, 1970, 75, 157-163.

Mahl, G. F. Disturbances and silences in the patient's speech in psychotherapy. Journal of Abnormal and Social Psychology, 1956, 53, 1-15.

Sarason, I. G. Characteristics of three measures of anxiety. Journal of Clinical Psychology, 1961, 17, 196-197.

Smith, E. K. Effect of the double-bind communication on the anxiety level of normals. Journal of Abnormal Psychology, 1976, 85, 356-363.

Wonderlic, E. F. Wonderlic Personnel Test, Form V. E. F. Wonderlic, P. O. Box 7, Northfield, Ill., 1959.

Zuckerman, M., & Lubin, B. Manual for the multiple affect check list. Educational and Industrial Testing Service. San Diego, California, (92107), 1965.

Footnotes

The authors wish to thank Mary E. Klein and Julie Klein for their technical assistance in preparing and coding the audio data and Roy Gabriel and Michael Thomas for their helpful reviews of earlier drafts. Correspondence should be sent to David M. Dush, Department of Psychology, Kent State University, Kent, Ohio, 44266.

TABLE 1.

Intercorrelations of Criteria and the Non-Ah Index[a]

	Non-Ah Ratio	*MAACL Depression*	*MAACL Anxiety*	*MAACL Hostility*	*Test Anxiety Scale*	*General Anxiety Scale*	*Debilitating Anxiety Scale*
Non-Ah Ratio		.038	.060	-.071	.043	.149	.130
MAACL Depression			.612**	.687**	.140	.083	.079
MAACL Anxiety				.668**	.319**	.208*	.185*
MAACL Hostility					.178*	.036	.015
Test Anxiety Scale						.633**	.703**
General Anxiety Scale							.555**
Debilitating Anxiety Scale							

[a]*The Non-Ah and MAACL correlations are based on recording 2. Correlations not involving the Mahl are based on all 98 subjects.*

*$p < .05$.

**$p < .001$.

TABLE 2.
Factor Analysis of the Individual Non-Ah Ratios[a]

Non-Ah Categories	*Communality Estimates*	*Factor I Standardized Loadings*	*Factor II Standardized Loadings*
Sentence Change	.669	.816	-.056
Repetition	.286	.532	.052
Stutter	.484	.688	.106
Omission	.589	.749	-.166
Sentence Incompletion	.655	.241	.772
Tongue Slip	.273	.210	-.478
Intruding Incoherence	.623	.009	.789
Eigenvalue		2.087	1.491
Per cent of Total Variance		29.8	21.3

[a]*Factors were derived from each subject's averaged scores from the three sets of Mahl scores, using a principal components analysis and orthogonal rotations.*

TABLE 3.
Canonical Correlation with the MAACL Criteria[a]

Variable Name	RECORDING 1	RECORDING 2		
	First Function Loadings	First Function Loadings	Second Function Loadings	Third Function Loadings
VARIABLES OF THE PREDICTOR SET				
Sentence Change	-.089	.384	-.466	.572
Repetition	.107	.595	-1.129	.073
Stutter	.034	-.180	.569	.108
Omission	.161	.314	.951	.144
Sentence Incompletion	-.075	-.159	.165	.955
Tongue Slip	1.100	.305	.713	-.592
Intruding Incoherence	-.140	-.486	-.311	-1.277
VARIABLES OF THE CRITERION SET				
Anxiety	.225	1.321	-.075	.027
Depression	-1.376	-.837	-.681	-1.292
Hostility	.681	.011	1.403	.438
Canonical Correlation	.710	.717	.443	.391
Probability	.001	.001	.009	.041

[a]*Loadings are standardized canonical function weights.*

FIGURE I

A. Calculation of the traditional "Non-Ah" Anxiety index: (Includes 7 Non-Ah disturbance categories)

$$\frac{\text{\# of Non-Ah Disturbances}}{\text{Total Words Spoken}}$$

B. Calculation of individual (Non-Ah) disturbance category ratios with the present "correction" to include "Ah's":

$$\frac{\text{\# of Disturbances in the Category}}{\text{Total Words} + (\text{Ah})^2 - (\text{Ah})^3}$$

FIGURE 2
AN ILLUSTRATION OF THE IMPACT OF THE *AH* CORRECTION FORMULA